Long-term tracking of social structure in groups of rats

Máté Nagy[1,2,3,4,5,6,†*], Jacob D. Davidson[4,5,6,†*], Gábor Vásárhelyi[1,3], Dániel Ábel[1], Enikő Kubinyi[7,8,9], Ahmed El Hady[4,5,6], and Tamás Vicsek[1,3]

[1]Department of Biological Physics, Eötvös Loránd University, Budapest, Hungary
[2]MTA-ELTE 'Lendület' Collective Behaviour Research Group, Hungarian Academy of Sciences, Budapest, Hungary
[3]MTA-ELTE Statistical and Biological Physics Research Group, Hungarian Academy of Sciences, Budapest, Hungary
[4]Department of Collective Behaviour, Max Planck Institute of Animal Behavior, Konstanz, Germany
[5]Department of Biology, University of Konstanz, Konstanz, Germany
[6]Centre for the Advanced Study of Collective Behaviour, University of Konstanz, Konstanz, Germany
[7]Department of Ethology, Eötvös Loránd University, Budapest, Hungary
[8]ELTE NAP Canine Brain Research Group, Budapest, Hungary
[9]MTA-ELTE Lendület 'Momentum' Companion Animal Research Group, Budapest, Hungary
[†]These authors contributed equally to this work
[*]nagymate@hal.elte.hu, jdavidson@ab.mpg.de

ABSTRACT

Rodents serve as an important model for examining both individual and collective behavior. Dominance within rodent social structures can determine access to critical resources, such as food and mating opportunities. Yet, many aspects of the intricate interplay between individual behaviors and the resulting group social hierarchy, especially its evolution over time, remain unexplored. In this study, we utilized an automated tracking system that continuously monitored groups of male rats for over 250 days to enable an in-depth analysis of individual behavior and the overarching group dynamic. We describe the evolution of social structures within a group and additionally investigate how past behaviors influence the emergence of new social hierarchies when group composition and experimental area changes. Notably, we find that conventional individual and pairwise tests exhibit a weak correlation with group behavior, highlighting their limited accuracy in predicting behavioral outcomes in a collective context. These results emphasize the context-dependence of social behavior as an emergent property of interactions within a group and highlight the need to measure and quantify social behavior in more naturalistic environments.

Introduction

Collective behavior emerges based on interactions between individuals in a group. This is observed at many different scales, from wound healing at the cellular level[1], to task allocation in social insects[2], group search behavior[3], and information exchange on human social networks[4]. Hierarchical structures are common in animal groups, for example, in the grooming relationships of chimpanzees[5], leadership and movement of pigeons[6], and reproduction in cichlid fish[7]. Social network structure influences decision-making[8,9], and dominance position within a network can influence an individual's fitness[10].

With rats, previous work has shown that individuals within a group have a social status related to dominance and that aggression and avoidance behavior are key elements of social interactions[11–13]. However, it is not known how individual interactions lead to the overall social structure of the group and how social structures change over time[14]. Fortunately, new automated tracking methods enable long-term tracking of individuals within social groups, providing a quantitative description of behavior and interactions[15]. For example, recent work has analyzed the ontogeny of collective behavior in zebrafish[16], lifetime behavioral differences in honeybees[17], and how genetic relatedness corresponds to group social structure in mice[18].

While it is acknowledged that long-term multi-modal characterizations are required to describe complex social behaviors, there are still a number of challenges[19]. An approach applicable to both lab and field conditions is to tabulate appropriate pairwise interaction information of animals in groups and use this to define measures of an individual's position in the social hierarchy. With mice, both aggressive and non-aggressive interactions have been used to define dominance hierarchies[20–26]. With primates, for example, pairwise "supplanting" interactions, such as when one individual displaces another from a food source, have been used to determine an individual's ranking[27–29].

Automated identification of approach-avoidance interactions has been used in previous work as a scalable method to characterize group hierarchy and leader-follower relationships[6]. Previous work with rodents, including rats[30] and mice[15], has

also considered approach and avoidance events. An approach-avoidance event occurs when one individual approaches another, but the other individual moves away (either by retreating or escaping).

There are multiple ways to compute social rankings based on pairwise interactions such as approach-avoidance events. The Elo score, which was originally developed to rank chess players and predict the outcome of future matches[31], is commonly used in animal behavior to calculate an individual's ranking using pairwise contest or interaction information[32–34]. If network "flow" is defined from winners to losers in an interaction network, the network measure of "flow hierarchy," which refers to how information flows through the network, can also be used to quantify hierarchical structure. The local reaching centrality (LRC) considers flow hierarchy in a network and quantifies a node's ability to efficiently reach other nodes in its immediate neighborhood[35]. Since "flow" occurs in the direction from winner to loser (or dominant to less-dominant), dominant individuals have higher local reaching centrality. Global reaching centrality (GRC) is calculated using the distribution of LRC scores, and this metric has been used to quantify the steepness of hierarchical structures in many different systems, including groups of horses[36], ant colonies[37], brain networks[38], industrial trade networks[39,40], and scientific citation networks[41]. In animal behavior, in addition to the individual Elo scores, metrics employed to measure social dominance hierarchies also include directional consistency of contests or interactions, proportions of interactions based on rank, transitivity, linearity of the hierarchy, and David's score or Elo score distributions across the group[24,42–45].

In this work, we developed an open-source vision-based automated tracking and behavioral characterization system to analyze the social behavior of small animals like rodents. We use this system to continuously monitor interactions and behavior of rat groups, enabling us to quantitatively examine the temporal evolution of social structure and the roles of individuals within these groups. We note, however, that rodent groups have inherently complex behavior and social structures, making it difficult to draw broad conclusions about the rules and processes that govern observed structures. We, therefore, focus on characterizing the results of our experiments and on using multiple metrics to describe various dimensions of the social structure.

For an extended 36-week period, we tracked the social behavior of 28 rats divided into several groups, and calculated behavioral metrics and interactions to analyze both individual and group behavior. We examined how individual behavioral differences persist and combine to form new social structures when the composition of the group is altered. At first, the rats were divided into 4 groups of 7, and following this, we merged groups. For the final sequence of group experiments, we then created 4 new groups of 7 and altered the size of the living areas. Following the completion of the group experiments, we ran individual behavioral assays on each rat and compared the results to those of the group experiments. This combination of methods and experiments enabled us to (1) identify a wide range of individual locomotion and social behaviors, (2) investigate the formation and details of dominance hierarchies, (3) investigate the effect of group composition changes and the associated social stress on the behavior of rats living in groups in enriched environments, and (4) compare behavioral assay results to behavior observed in a group setting. Overall, our work demonstrates scalable methods for describing long-term changes in animal group social structure and emphasizes the need to use such methods to obtain a full picture of group social structure and interactions in natural or semi-natural environments.

Results

Long-term tracking and quantifying individual behavior

We tagged individuals with color markers and employed automated tracking to determine each rat's movement over time (Figure 1A-D). Over the course of the experiment, we performed manipulations to alter the group composition and the living area available for the group to use. We used two different breeding lines of Wistar laboratory rats, denoted A and B (see Methods for details), with associated individual labels $a*$ or $\alpha*$ and $b*$ or $\beta*$, respectively (see Figure 1E). We initially divided the rats into four groups of seven, with A rats in groups $A1$ and $A2$ and B rats in groups $B1$ and $B2$. The rats remained in these groups for the first observation period, which lasted a total of 21 weeks - we denote this time as phase 1. Following this, in phase 2, we merged groups $A1$-$A2$ and $B1$-$B2$ for three weeks and then merged all for three weeks by opening portals between their compartments. For the final series of group experiments in phase 3, rats from each original group were mixed together to create four new groups. The reshuffling in phase 3 was done according to body mass at the end of phase 2 (mean = 480 g; min = 364 g, max = 613 g; Q1 = 423 g, median = 481 g, Q3 = 532 g) , allocating rats to new groups by ensuring that each group had the full range of masses and included members from every previous group ($G1_{min}$ = 394 g, $G1_{max}$ = 541 g, $G2_{min}$ = 372 g, $G2_{max}$ = 587 g, $G3_{min}$ = 400 g, $G3_{max}$ = 613 g, $G4_{min}$ = 364 g, $G4_{max}$ = 555 g). Figure 1E shows the experimental structure and the associated measurement periods.

We calculated automated trajectory-based behavioral metrics to quantify behavior over the duration of the experiment. We calculated and averaged each metric over successive time periods of 3 weeks (denoted as Pd), with associated numbers 1-12. We use the summary metrics to ask how behavior changes over time, how individuals differ, how groups differ, and how previous individual and group behavior predicts changes when new groups are formed.

To assess dominance-related interactions and social structure, we tabulated approach-avoidance events between all pairs of rodents in each group. This automated method defines "events" as when a pair of rats come close to each other: the "displacer",

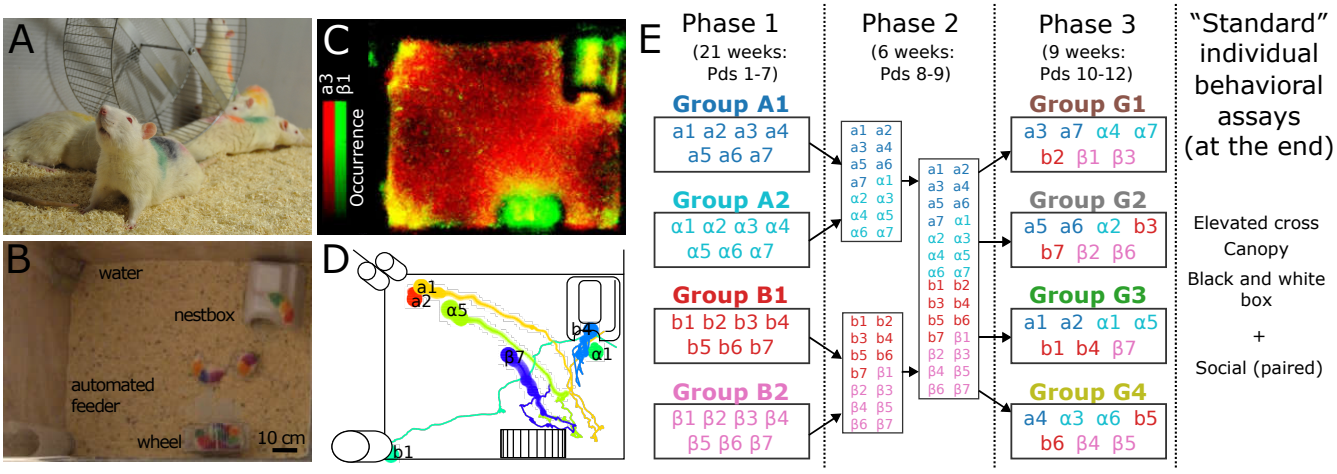

Figure 1. Experiment setup and timeline. (A) Photo of the rats with color-codes for individual identification and tracking. (B) Still image from the video that was used for tracking (from group $G1$, during Pd 10) taken by a light-sensitive camera at low lighting conditions. Image overlaid with labels indicating the important objects (water, nestbox, etc.). (C) Continuous tracking allowed for the reconstruction of each individual's space use. The heatmap shows the space use of two rats during a 3-week period at the beginning of phase 3. Areas used only by $a3$ are shown with red, only by $\beta 1$ with green, and areas visited by both (e.g. at the water and the feeder) are shown with yellow. (D) Trajectories were used to identify dominance interactions in the form of approach-avoidance events, where one individual approaches another, but the other moves away (by backing up or fleeing). Shown is an example of trajectories from group $G3$ in period 10. Lines show locations for 60 seconds, with the semitransparent circles of increasing size showing the more recent positions. (E) Overview of experimental manipulations. We calculate behavioral metrics over each 3-week "period" (abbreviated as Pd). Phase 1 had rats in original breeding line-sorted groups $A1$-$A2$ (line A), $B1$-$B2$ (line B), for a total of 7 periods. Each rat is labeled with lowercase letters a/α or b/β according to breeding line. Individual numbers within each group are sorted in ascending order according to rank as determined by Elo score at the end of phase 1, i.e. $a1/a7$ were the highest/lowest ranking individuals in $A1$ during Pd 7, $\alpha 1/\alpha 7$ were the highest/lowest in $A2$, etc. In phase 2, the groups were mixed together by breeding line during Pd 8, and then all together for Pd 9. At the beginning of phase 3 (Pd 10), new groups were formed ($G1$-4). During Pds 11 and 12 in phase 3, the compartment area sizes were changed (see Methods and Fig S8). At the end of the experiments, individual behavior was assessed by traditional individual and pairwise assays.

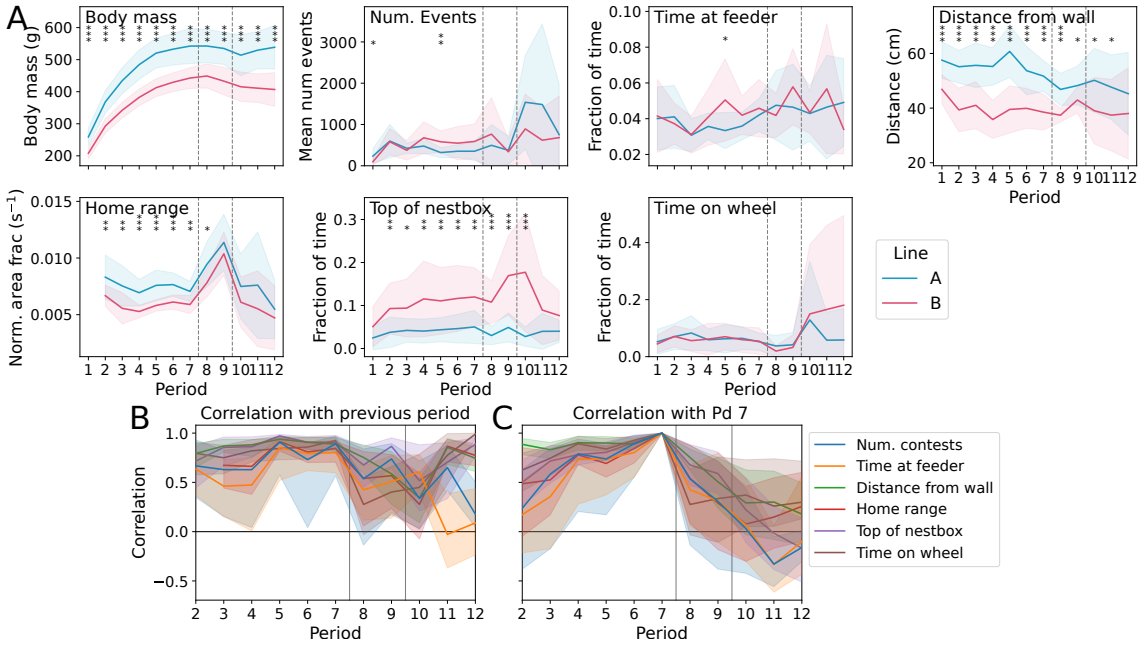

Figure 2. Breeding line comparison and correlation. (A) Per-line body mass, average number of events, and space use metrics. Significant differences between breeding lines for a designated period, as determined with a T-test for difference in means, are denoted as follows: $p < 0.05$ with *, $p < 0.01$ with ** and $p < 0.001$ with ***. See also Fig S3 for space use compared according to group, and Fig S4) for space use metrics for each individual rat. (B) Correlation with the previous period, calculated across all rats with respect to a particular metric. Shaded area shows confidence interval calculated via bootstrapping. Note that values significantly different from zero are when the confidence intervals do not contain zero. (C) Correlation with Pd 7 (the last measurement in phase 1). Shaded area shows confidence interval calculated via bootstrapping.

i.e. the dominant rat in an event, subsequently stays in place or continues moving forward, while the other (the "displaced", i.e. subordinate rat) move away[6]. This type of approach-avoidance interaction can also be dynamic, such as when one individual chases another. We use the matrix of approach-avoidance events to calculate metrics that describe the dominance structure of each group and each individual's position in this structure.

Breeding line and group differences

We use automated measures of space use and pairwise interaction events to characterize individual and group behavior. We first examine general differences between breeding lines.

In the beginning, the rats were juveniles and were growing rapidly, as shown by the large increases in body mass during this time period. The *A* rats were, on average, significantly larger than the *B* rats during each period (T-test comparing average mass of *A* rats to *B* rats yields $p < 0.001$ for each period). All rats had approach-avoidance events during the experiment, and there were no consistent significant differences among the breeding lines. However, there was an increase in the number of events per rat in phase 3 compared to phases 1 and 2 (Mean number of events per rat in phases 1, 2, 3, respectively: 447, 494, 976; T-test mean of phase 1 to phase 2, p=0.64; mean of phase 1 to phase 3, p=0.0044; mean of phase 2 to phase 3: p=0.0176).

The metrics of time at feeder, distance from wall, home range[46], time at top of nestbox, and time on wheel describe space use. While the breeding lines did not have general differences in time at feeder or time on wheel, line *A* rats tended to be farther from the wall, visited more parts of the living compartment (larger home range), and spent less time on top of the nestbox in comparison with *B* rat groups. However, while these differences were clear during phase 1, the differences in distance from wall and home range decreased when the lines were mixed, with home range no longer significantly different from Pd 9 onward, and distance from wall no longer significant in Pd 12. Breeding line differences in time spent at top of nestbox showed a large increase when group membership was changed in Pds 9 and 10, but subsequently decreased and were not significant in Pds 11 and 12 (Fig 2). Note that one group (G1) in phase 3 displayed a different pattern of wheel usage than other groups, with several rats spending a very large amount of time on the wheel at the same time and thus unable to use it for running (Figs S3 and S4); however, there were no breeding line differences in this behavior. Overall these metrics suggest that the different breeding lines differed in their space use tendencies, but differences decreased when rats were placed in mixed groups in phase 3.

We quantify changes in individual behavior using the correlation coefficient across periods. This shows that individuals

have consistency in number of events and space use, as demonstrated by the generally positive correlations during the entire observation period (Fig 2B). However, while there is consistency from one period to the next, Fig 2C shows that small behavioral shifts over time can accumulate. Moreover, we see that the re-groupings facilitated changes in behavior. This is demonstrated by the sharper decrease in the correlation of behavioral metrics with Pd 7 in phases 2 and 3 compared to that in phase 1 preceding Pd 7. In particular, while the correlation coefficient for home range and time at top of nestbox during Pds 11 and 12 showed high correlations (Fig 2A), the correlation of these 3 measurements with Pd 7 values was lower (Fig 2B). For example, for Pd 12, the correlation with the previous period for home range was 0.77 (95% CI: [0.68 0.87] and for top of nestbox was 0.89 (95% CI: [0.71 0.95]), while the correlation values with Pd 7 were 0.26 (95% CI: [-0.13 0.57]) and -0.17 (95% CI: [-0.5 0.16]), respectively. This indicates that the new behavioral routines of phase 3 differed from those of phase 1.

Metrics for group social structures

With the pairwise approach-avoidance interaction matrices for each period, we use multiple metrics to characterize different aspects of group social structure and an individual's placement in this structure. The metrics to characterize individual social placement include Elo score, David's score, local reaching centrality, and fraction of events dominated, and those to characterize group social structure include Elo score steepness, David's score steepness, global reaching centrality, directional consistency index, and triangle transitivity index. In this section we use idealized networks (shown in Fig 3) to illustrate what the group social structure metrics represent. Note that while other work has used similar idealized or artificial networks as "categories" to label group social structure[47], here we use the ideal networks (including connected hierarchy, line, layered hierarchy, layered-half, non-transitive, single dominant, single out, and symmetric) not as categories, but rather to give intuition for how the different metrics describe different aspects of the social structure. In the following section, we report the metrics for each group and use them to describe the experimentally observed structures.

The Elo score steepness (ESS) is a measure of the spread of the distribution of Elo scores across the group. It is calculated by converting the Elo score to a success probability, summing normalized values across group members, and calculating the slope of a linear regression fit to the resulting values[45]. The David's score steepness (DSS) (often referred to simply as hierarchy 'steepness', or 'classic steepness'[42,45]) is calculated as the slope of a linear regression fit to the normalized David's scores among group members[48]. Individual local reaching centrality (LRC) uses the directed network of excess pairwise event outcomes (positive entries for rats in a pair that was dominant in more events, and zero for the other rat - see Methods) in order to assign higher scores to individuals in higher positions within a group hierarchy. For an unweighted directed network, LRC is the fraction of nodes reachable by any given node; a generalization of the metric accounts for weighted connections[35]. Global reaching centrality (GRC) is the average difference of nodal LRC with that of the highest LRC of any node in the graph, and a higher GRC indicates a more hierarchical network[35].

The directional consistency index (DCI) is the fraction of events dominated by the more dominant individual of each pair, with 1 corresponding to perfect predictability in the outcome of a pairwise event (i.e. one individual is always dominant), and 0 representing an exchange of approach-avoidance outcomes (i.e. each individual dominates the same number of events)[42,49]. The triangle transitivity index (TTRI) is the fraction of triad relationships that show transitivity in pairwise event dominance outcomes (i.e. if $a \rightarrow b$ and $b \rightarrow c$, then $a \rightarrow c$ for a transitive triad)[43,50].

From Fig 3 we note that the ESS and DSS, which both aim to measure the steepness of hierarchy within a group, show similar trends at times and differ at others; both have high values for the connected hierarchy network but differ for the line network. We also note that the aspects of the network structure described by ESS/DSS versus GRC are different (c.f. differences in the connected hierarchy, layered-half, and single dominant networks); the former is maximized when a well-connected structure exists (i.e. the hierarchy shows a clear distribution that lends itself to a linear regression fit), while the latter is maximized when more extremes in hierarchical structures exist (for example, the single dominant). Although a comprehensive evaluation of these metrics is beyond the scope of this study (see, for example,[45]), here we calculate and examine multiple metrics to ensure a robust interpretation of the data, as well as to facilitate comparison of our findings with other assessments of group social structure found in the literature.

Group social structures

We find that groups differ in their social structure, but within-group structure shows consistency when group membership remains unchanged. We compare phase 1 and phase 3 group social structures because the associated periods all featured groups of 7 rats. We show results for all group social structure metrics, as well as the mean number of events and fraction of events with the dominant rat, and note instances where the trends for the metrics are similar versus contrasting.

In phase 1, in contrast to space use, which showed clear breeding line-based differences (Fig 2), we do not see clear differences between lines *A* and *B* in terms of overall group social structure. While the fraction of events with the dominant rat was higher for the *A* groups in comparison to the *B* groups, other metrics do not show large or consistent differences (Fig 4). Within phase 1, each group showed consistency in the social structure over time, with the exception of Pd 6 for group *A*2,

	Connected hierarchy	Line	Layered hierarchy	Layered-half	Non-transitive	Single dominant	Single out	Symmetric
ESS	**1.00**	**1.00**	0.81	0.79	0.65	0.59	0.59	*0.55*
DSS	**0.99**	0.08	0.41	0.32	0.12	0.11	0.11	*0.00*
GRC	0.58	0.58	0.78	**1.00**	0.00	**1.00**	*0.03*	-
DCI	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	*0.00*
TTRI	**1.00**	-	**1.00**	**1.00**	*0.00*	-	-	-

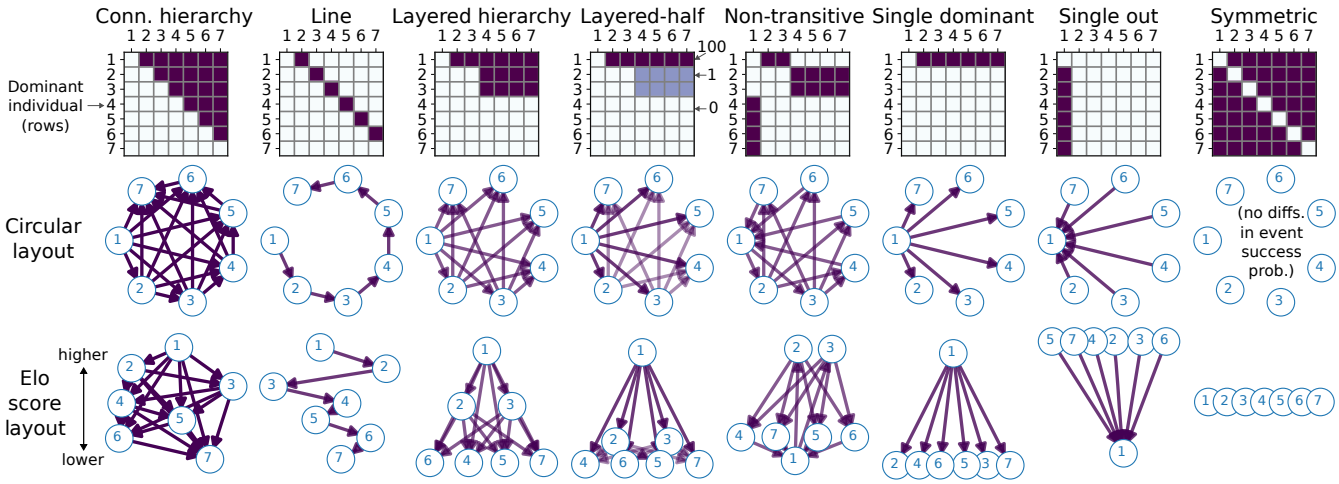

Figure 3. Idealized networks and group social metrics. The table at the top shows the scores calculated: Elo score steepness (ESS), David's score steepness (DSS), global reaching centrality (GRC), directional consistency index (DCI), and triangle transitivity index (TTRI). Note that the GRC is not defined for the symmetric network, and the TTRI is not defined for networks that do not contain any dominance triads. The different idealized networks have 7 nodes, and individual entries are either 100 or 0 (for the layered-half network, 100, 1, and 0 are used). The connected hierarchy network has a non-symmetric structure. The line network has a single "line" of pairwise interactions, where each individual only interacts with one other. The layered hierarchy network has a single individual who dominates all others and two other sub-dominant individuals who only dominate the four others below them. The layered-half network has the same structure but lower values for the subordinate individuals. The non-transitive network has individual 1 dominating 2-3, 2-3 dominating 4-7, but 4-7 dominating 1.The single dominant network only has events with the dominant individual. The single-out network only has events with the subordinate individual. The symmetric network has equal event dominance probability among all pairs. In the table, the highest score for each metric (rows) is in bold, and the lowest score is in italics. The ESS is highest for the connected hierarchy and line networks, the DSS is highest for the connected hierarchy network while low for the line network, and the GRC is highest for the layered-half (i.e. structured but nonlinear hierarchy) and single dominant networks. For the symmetric network, the DCI is 0 because there are no consistent dominance relationships; for other networks, dominance is one-sided and the DCI is 1. The TTRI is 0 for the non-transitive network, and 1 for other networks where dominance triads are predicted. In addition to matrix plots, each network is visualized by showing connections in the direction of the more to less dominant individuals in each pair (note that no connections are shown for the symmetric network because, in this case, there are no differences in event dominance probability). We show both a circular layout and a layout based on Elo scores, where individual nodes with higher Elo scores are shown higher up on the *y*-axis.

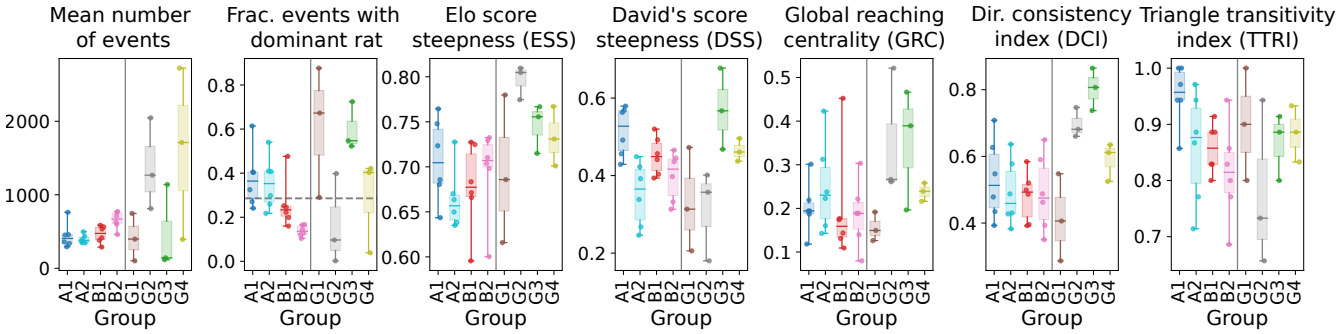

Figure 4. Measures of group social structure. The mean number of pairwise events and the fraction of total events with the most dominant rat are shown in addition to the hierarchy-related metrics of Elo score steepness (ESS), David's score steepness (DSS), global reaching centrality (GRC), directional consistency index (DCI), and triangle transitivity index (TTRI). The fraction of events with the dominant rat (where "dominant" is defined as the individual with the highest Elo score) is analogous to the measure of "despotism" used in other work[22]; the dashed line shows the expected value if all pairs of rats have the same number of events. The metrics are calculated for each period and are shown as boxplots for each phase 1 and phase 3 group. See also Fig S7 for values for each group over time.

where a large change in the individual rank ordering in the social structure of the group took place. This is seen in the network visualizations, as well as in the positive correlation of Elo scores from one period to the next (Fig 5A-B).

In general, we saw larger differences between the groups in phase 3 in comparison to those in phase 1. In phase 3, $G1$ and $G3$ each had a single consistent dominant individual, $G2$ had ongoing changes in social structure, and $G4$ had a stable hierarchy but with ongoing events. All groups had consistency in structure, but the correlation of individual scores with the groups was higher for groups $G1$, $G3$, and $G4$ in comparison with group $G2$. (Fig 5C-D).

Compared to other phase 3 groups, $G1$ and $G3$ had a relatively low number of events and a high fraction of events with the dominant individual. Each of these groups had a single individual that was consistently ranked as most dominant (see Fig 5C). However, in comparison to $G1$, $G3$ had on average higher DSS, ESS, and GRC. This and the higher DCI index suggests that group $G3$ had a steeper hierarchical structure than group $G1$.

In contrast to groups $G1$ and $G3$, group $G2$ did not have a single individual that remained dominant during each period. Group $G2$ had many events, the lowest fraction of events with the dominant rat, and the lowest transitivity (TTRI) of the groups. This and the lower correlation coefficient in Elo scores compared to other phase 3 groups suggest an ongoing struggle for position within the social network where ongoing events maintained pairwise relationships. However, we note the differences obtained in the hierarchy steepness measures for group $G2$: the David's score steepness suggests a weak hierarchy, while the Elo score steepness and GRC suggest hierarchies definitely exist.

Group $G4$ had similar patterns of metrics to Group $G3$, but with several distinct differences: these include lower magnitudes of ESS, DSS, GRC, and DCI, a lower fraction of events with the dominant individual, and overall many more events (although the mean number of events decreased dramatically from Pd 11 to Pd 12 – see Fig S7). With this, we can describe $G4$ as having a middle-steep hierarchy that was maintained by many ongoing events among pairs. This differs from $G1$ and $G3$, where the high fraction of events with the dominant individual suggests that the hierarchy was maintained mostly by these events.

During phase 3, the area available to each group was changed during Pds 11 and 12 by moving the compartment borders that separated the groups. In Pd 11, $G1$&$G4$ had a larger area and $G2$&$G3$ a smaller area. Pd 12 had these sizes switched. These manipulations did not have a consistent effect on space use or group social structure metrics (Fig S8).

Individual social rankings, changes over time, and body mass

We found that previous social status in phase 1 did not predict an individual's placement in the new group social structures of phase 3 (Fig 6A). This result holds if instead of using absolute Elo score values as shown in Fig 6A, rank scores of subordinate and dominant are used for the lowest two and highest two Elo scores, with other assigned as middle ranking (see Fig S6). Because the individual ranking metrics are correlated (Fig S5), for clarity we focus on showing results the often-used Elo score[32–34]; however, we also note that the other individual social metrics (including faction of events dominated, David's score, and local reaching centrality) showed similar trends (Fig S6) with respect to predicting individual placement.

Because social ranking can be used to regulate access to resources such as food, we further examine the relationship between social rank and weight gain/loss. We compare average individual social rankings to weight gain or loss during phase 3. At the beginning of phase 1, all rats were young and gaining weight. However, by the end of phase 1, the average weight gain from the previous period was small, and not all rats were still gaining weight. When the groups were merged in phase 2, the

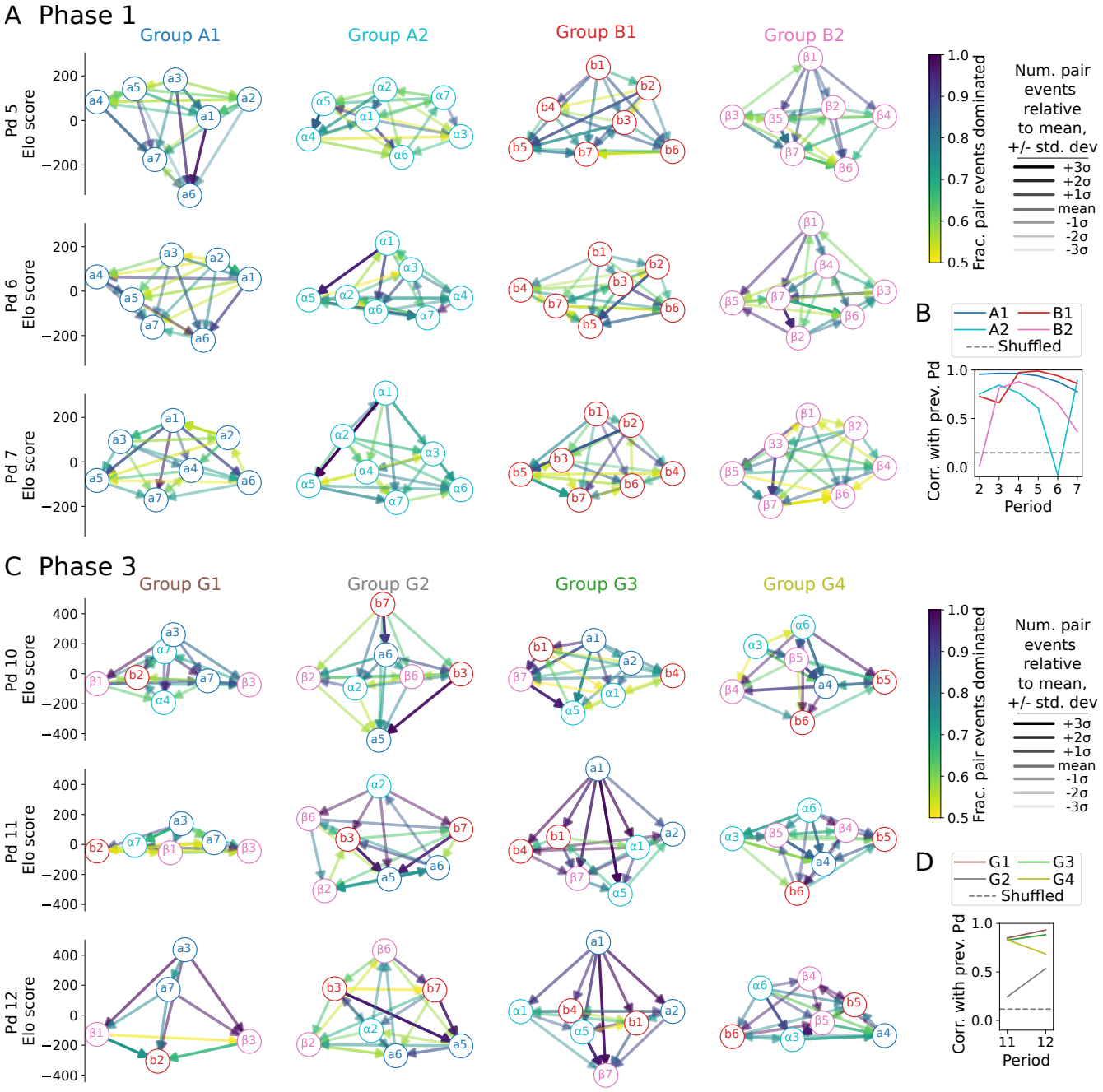

Figure 5. Social structure network visualizations. (A,C) A visualization of group networks during (A) the last 3 periods of phase 1, and (C) phase 3. Columns correspond to different groups and rows for each period. The position of each individual on the y-axis is set according to their Elo score. The direction of each connection indicates which individual dominated more events in the pair (e.g. a connection $a4 \rightarrow a6$ indicates that $a4$ more often displaced $a6$ than vice versa), the color indicates the fraction of events dominated, and the transparency is proportional to the total number of pairwise events relative to the mean for that group and period. (B,D) The correlation of individual Elo scores with the previous period, for (B) phase 1 groups, and (D) phase 3 groups. Dashed line shows baseline correlation value calculated by shuffling groups and periods during phase 1 (B) or phase 3 (D). See also Fig S5 for individual metrics (including num. events, fraction dominated, Elo score, David's score, and reaching centrality) for each individual rat plotted for each period.

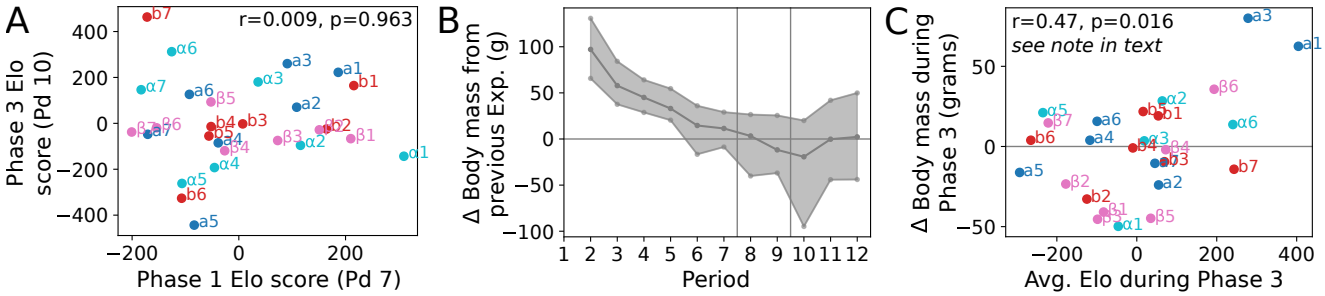

Figure 6. Phase 1 to phase 3 individual rankings, body mass changes and social ranking. (A) Comparison of individual rat ranking metrics at the end of phase 1 (Pd 7) to those at the start of phase 3 (Pd 10), after the new groups were formed. (B) Distribution of body mass changes over time for all individuals. The middle line is the mean, and rugged curves indicate the maximum/minimum across all individuals. (C) Average Elo score during phase 3 compared to body mass change during phase 3 for each rat. Change in body mass during phase 3 is calculated as the body mass in Pd 12 minus body mass in Pd 10.

average change in body mass (Δ body mass) continued to decrease and was negative for the last period of phase 2 and the first period of phase 3. Specifically, in the new groups of phase 3, the variance in the distribution of Δ body mass increased, with one rat (rat $\alpha 4$, which was subsequently permanently excluded from the experiment) losing nearly 100g relative to the previous period (Fig 6B).

Fig 6C shows that although the two most dominant rats during phase 3 gained the most weight (rats $a3$ and $a1$ from groups $G1$ and $G3$, respectively), average social dominance rank was not a robust general predictor of body mass changes across all rats. Including these dominant rats, the relationship between the average Elo score during phase 3 and the change in body mass during phase 3 has a significant correlation with the values shown in Fig 6C. However, this result is not robust: if these two rats are removed, the correlation drops greatly to $r = 0.16$ and is no longer significant ($p = 0.44$). This result also holds if subordinate-middle-dominant ranks are used instead of absolute Elo score values ($r = 0.48$ and $p = 0.013$ including all individuals; $r = 0.26$ and $p = 0.23$ removing rats $a3$ and $a1$). This demonstrates the complex relationship between dominance and body size in rodent social groups[51,52]. While there was likely a feedback regarding social rank and body mass for the two dominant rats in groups G1 and G3, respectively, it is difficult to link weight gain/loss to social rank in a general sense.

Individual metrics compared to behavioral assays

We used individual and pairwise assays performed after the group experiments to test the behavior of each rat. The individual black and white box, canopy, and elevated plus-maze results were used to define a composite boldness score. A pairwise social test with an unfamiliar rat, where two individuals are placed together and various behaviors characterizing interactions, such as sniffing the other, are scored (see Methods; Figure S10), was used to define a social interaction score. This social test, which is also referred to in the literature as the "reciprocal social interaction test", has been widely employed for behavioral phenotyping related to anxiety and autism[53–56].

In general, we find low and/or inconsistent relationships between behavior in groups and behavior in the assays (Figure 7A). This is shown by the comparison of behavioral metrics from the last period in phase 3 with the individual boldness and social interaction scores. In particular, the social metrics measured in a group setting, including the number of events and the Elo score, do not exhibit consistent or significant correlations with the pairwise social interaction score. Although we see positive correlations for the boldness score compared to the related metrics of distance from wall and home range, and a negative correlation with top of nestbox, these correlations are not significant ($p > 0.05$), with a notable remark that the 2 most dominant individuals ($a3$ of $G1$, $a1$ of $G3$) have high boldness scores within their group. However, when considering the individual assays separately, we do find a significant correlation between top of nestbox and time spent in the open area during the elevated cross assay (Fig S11A). We also find that breeding line and group membership do not consistently predict differences in individual test scores (Figure 7B).

The social interaction score shown in Figure 7A-B was obtained via pairwise behavioral assays performed with an unfamiliar rat. We repeated these tests with a second unfamiliar rat in order to test repeatability (tests were also performed with a familiar rat from the respective phase 3 groups - see Figure S13). The composite scores from the tests with the second unfamiliar individual show a low correlation with the scores obtained with the first unfamiliar individual (Figure 7C, $p > 0.05$).

Other work has noted that individual behavior in assays may depend not only on an individual's social dominance status, but also on the nature of the social hierarchy of the group to which the individual previously belonged[47]. To test this, we fit a linear regression model predicting individual behavior assay results based on a combination of individual metrics and the Elo score steepness (ESS) of the group where the individual was located during Pd 12. While including this additional information

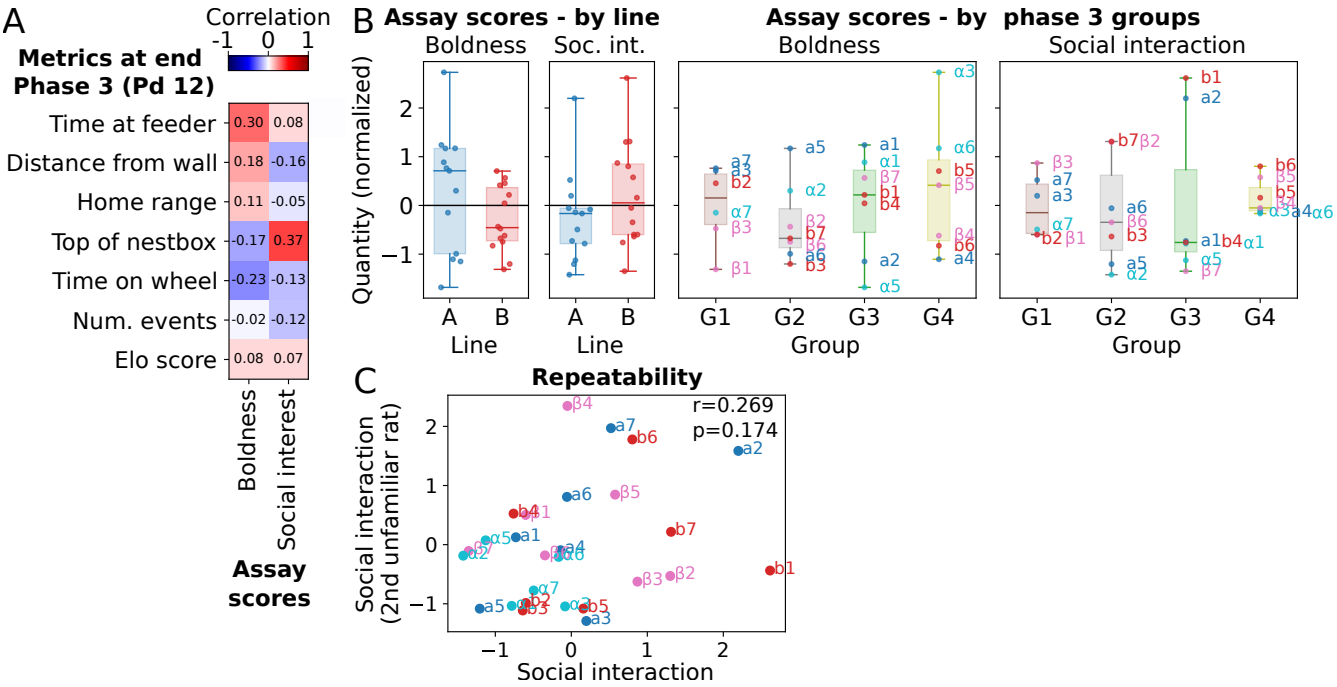

Figure 7. Behavioral metrics at the end of phase 3 compared to assays. (A) Pearson correlation values for space use and social behavioral metrics from the final period in phase 3 (Pd 12) with individual assay scores. Labels and color scales denote correlation values. Note that none of the correlations are significant (all p-values > 0.05, calculated using t-distribution). (B) Individual score distributions according to breeding line (left), and by phase 3 group membership (right). Scores are normalized by the mean and standard deviation of values measured for all rats. (C) Comparison of social interaction scores calculated from tests with a first unfamiliar rat (x-axes; values shown in A-B), with scores calculated from tests with a second unfamiliar rat (y-axes). See also Figures S11 and S13.

increased explanatory power, we did not observe consistent significant patterns (Fig S12).

Discussion

We utilized automated tracking techniques to describe how rat groups develop and maintain a dynamic social structure over time, as well as how the social structure changes after regrouping. Across successive periods, we observed a general consistency in the behavior of both individuals and groups. However, considering longer periods of time across multiple periods, we observed that the gradual accumulation of small changes can result in substantial behavioral changes over these longer time scales. In addition, when the group composition was altered, we observed accelerated changes in behavior. Multiple metrics, including the Elo score, David's score, and reaching centrality, were employed to describe the overall hierarchical structure of each group as well as individual social rankings; these metrics revealed both similar and dissimilar aspects of the structure. Different groups can vary a lot in their structure, and in particular, we found different and distinct social structures in the newly formed groups of phase 3. We found that an individual's new position in the social hierarchy cannot be predicted based on their prior status when group composition changes. While conventional individual assays (including the elevated cross, canopy, and other tests) produce consistent test results, we found that these measures have little correlation with individual behavior in a group setting. Moreover, we found low repeatability in the scores measured with standard social test assays by performing the same test with different individuals; this also contributes to why behavior assays have little correlation with group behavior.

At the beginning of phase 1, the rats were still juveniles. Social interactions, particularly those related to aggression and dominance, are known to develop over time[11,57,58]. Our observations are consistent with this, in particular, because we observed an increase in the number of approach-avoidance events and fights at the end of phase 1 (Figure S2). It is likely that in addition to group composition and interactions, the development stage also influenced the number of events and the differences in social structure from phase 1 to phase 3. In a natural population, groups consist of individuals of different ages[57]. Targeted group mixing experiments – for example, with both juveniles and adults in the same group – could be used to ask how these effects interplay to generate overall emergent group social structures.

Behavioral assays are often used to quantify the behavior of rodents, and many new tools are being developed for both

individual and social tracking and behavioral scoring[59–64]. Tests that have been used to quantify social behavior in rodents include, for example, reciprocal social interaction, social approach partition, social preference, social transmission of food preference, food allocation, and reciprocal cooperation[13,54,65]. Most of these tests rely on creating artificial situations in order to measure the corresponding behavioral outcome. We compared social behavior measured in a group setting to scores of individuals measured with the reciprocal social interaction test. The pairwise social interaction test has known limitations, such as environmental dependencies, the possibility of aggression, and limited clarity as to which rat-initiated interactions[66,67]. Although other behavior assays, such as the social preference test using a T-maze or modified home-cage tests, attempt to address limitations, these assays also have their own limitations[66,67]. It is an open question as to whether the social behavior measured in such tests can predict the social behavior under more natural conditions that include complex social interactions[68]. In this respect, social interest/interaction and social dominance may represent different aspects of behavior, with the latter possibly only able to be measured in a group setting. Our comparison highlights the need for further work in this area.

We also note that while our group-living experiments provided space for complex environmental and social interactions, the conditions in the experiments were still much different from those that rats experience in the wild. In this respect, our methods are similar to recent studies with mice, such as experiments with groups in the "social box"[20–22]. These experimental setups can, therefore, be described as semi-natural. A unique aspect of our study is the extended observation period, which allowed us to examine not only group social structure but also its change over time.

The study of social behavior is particularly important in animal models utilized for the understanding and treatment of social-related neuropsychiatric disorders. Rodents such as mice and rats have been indispensable as model biological organisms, with particular relevance to clinical research due to their short lifespan and tolerance to laboratory environments[69–71]. However, the laboratory environment typically restricts the development of complex social behaviors; for example, rats are often kept in small cages and then transferred to separate environments to examine social interactions using simplified behavioral assays like the 3-chambered social preference test[56]. It is therefore not surprising that the translational relevance of these tests is limited[72,73]. Incorporating environmental and social complexity into experiments can increase the generalizability of conclusions drawn from laboratory studies[74,75]. Moreover, long-term studies to examine group behavior may be an essential component to include in translational research applications, for example, in the testing of psychotherapeutic drugs to treat social anxiety[72,73]. In particular, an important topic for future work is to establish standardized and reproducible tests and measures that are properly representative of a full range of social behavior[75,76]. Furthermore, we note that much of our neuroscientific understanding of social behavior comes from dyadic interactions and reduced forms of social interactions[74,77–79]. In light of the growing interest in the neuroscience of natural social behavior[68], going beyond basic social testing paradigms lends the opportunity to unravel a richer repertoire of neural mechanisms.

We note that while our social behavior analysis was used with video tracking data, it could be applied to other types of data, for example, data derived from markerless tracking methods, motion capture or QR-code tracking. Future work can continue to expand on methods in this area, for example, including more detailed posture data, which can be used to describe social interactions in more detail[80–83]. Systems in this direction have already been developed for use with mice[61]. While we used automated detection of approach-avoidance interactions to define pairwise events, we note that a more detailed approach could use a combination of multiple behavioral interactions in order to define event "winners" and "losers"[84]. Moreover, detailed insight could be gained by using a hybrid method when the automated detection is followed by manual scoring of the behavior [85,86]. Approach-avoidance as a measure of dominance has limitations, as the subordinate animal can freeze in place, with lack of movement signifying its subordinate status[87] and this would be missed by the automated scoring scheme. Another area for future work is testing the functional consequences of group composition and social structure on individual or group performance, for example, with respect to search[3].

Author contributions

T.V. and M.N. conceived the idea and designed the project. M.N., G.V., D.Á., and E.K. conducted the experiments and collected the data. E.K. performed the analysis of the individual assays. G.V. designed and wrote the software framework for tracking the individuals. M.N., G.V., and V.T. designed and performed the analysis of the behavior within the groups. J.D.D., M.N., and A.E.H designed the analysis of the group-level data. J.D.D. performed the analysis of the group-level data. J.D.D., M.N., and A.E.H wrote the initial draft of the manuscript, and all authors revised the manuscript.

Acknowledgments

We thank all the people who helped during the experimental work, especially Valéria Németh, Gergő Somorjai, Zsuzsa Ákos, Katalin Schlett, Péter Urtz, and András Péter (for Solomon coder). We acknowledge funding from Eötvös Loránd University and the Max Planck Institute of Animal Behavior. This research was partially supported by grants to T.V. establishing the MTA-ELTE Statistical and Biological Research Group, including NKFI Office Grant OTKA SNN 139598. Other sources of

funding include the Hungarian Academy of Sciences (a grant to the MTA-ELTE ‘Lendület’ Collective Behaviour Research Group, grant number 95152, National Brain Programme 3.0 (NAP2022-I-3/2022), and the Companion Animal Research Group (PH1404/21), and the DFG Centre of Excellence 2117 “Centre for the Advanced Study of Collective Behaviour” (ID: 422037984). M.N. was partially supported by the Hungarian National Research, Development and Innovation Office (grant no. K 128780) and by Isaac Newton Institute for Mathematical Sciences, University of Cambridge. J.D. was partially supported by the Heidelberg Academy of Sciences Young Scientist (WIN-Kolleg) program.

Methods

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Mate Nagy (nagymate@hal.elte.hu).

Materials availability

The study did not generate new unique reagents.

Data and code availability

Trajectory and behavioral metrics data are available through Zenodo at doi.org/10.5281/zenodo.7615468. Recorded video sequences were analyzed offline with a custom-written software to obtain individual positions and orientations (source code available at github.com/vasarhelyi/ratognize), as well as trajectories and metrics (source code available at github.com/vasarhelyi/trajognize). For details, see SI Methods of[3]. The scripts to run analyses included in this paper are available at github.com/jacobdavidson/ratsocialgroups.

Experimental model and subject details

Subjects

We used 28 Wistar male rats from 2 inbred breeding lines (14 Crl:WI BR and 14 HsdBrlHan:WIST, 7 litters/line, 2 individuals/litter; ordered from Toxi-Coop Zrt, Hungary) in this study. The rats arrived on 24 May 2011 at an age of 6 weeks. Rats were separated into 4 same-line groups (i.e. the phase 1 groups), each containing 7 rats from different litters, and were housed and tested together.

Each rat was marked with a unique 3-color barcode on its back using nontoxic "Special Effects" hair dye in 5 distinctive colors (Red: Nuclear Red, Orange: Napalm, Green: Sonic Green, Blue: Londa color 0/88, Purple: 4 units Atomic Pink and 1 unit Wild Flower). These codes were applied/renewed every 3 weeks.

Ethical guidelines

The procedures comply with national and EU legislation and institutional guidelines. The experiments were performed in the animal facilities of Eötvös University, Hungary, and in accordance with Hungarian legislation and the corresponding definition by law (1998. évi XXVIII. Törvény 3. §/9. — the Animal Protection Act), which states that noninvasive studies on animals bred for research are allowed to be performed without the requirement of any special permission (PE/EA/1360-5/2018).

Method details

Experimental conditions and monitoring

Animals were housed in 4 compartments (sized 100 x 125 x 100 cm^3) with polypropylene covered wooden walls and sawdust changed weekly on a tiled floor (see Fig S1). Their room was kept at a controlled temperature of 21 ± 2°C, and with controlled light conditions featuring a daily cycle with 13h/11h dark/light. The dark (active) period was from 6am to 7pm with illumination at floor level ∼3-4 lux; the light period followed from 7pm to 6am of the following day with illumination of 300 lux. We video recorded the compartment 24/7 using a low-light sensitive camera fixed to the ceiling (Sony HDR-AX2000, 2.9 $\times$ 1.8 m^2 field of view, 1920 $\times$ 1080 resolution, 25 fps de-interlaced). Rats had *ad libitum* access to water and a shelter (nestbox), and access to food based on a fixed schedule for automated feeding. The feeding schedule followed a weekly cycle: 3 days (Sat, Sun, Mon) access to food 3 times for 1 hour (at 6am, noon, and 6pm); 3 days (Tue, Wed, Thu) access to food 2 times for 1 hour (at 6am and 6pm); and 1 day (Fri) access to food *ad libitum* between 6am and 7pm). The housing compartments were cleaned once a week.

We measured the weight of each individual three times a week (Mon, Thu, Fri; between 5pm and 6pm) and inspected their health. Some rats had injuries, and When a rat had a larger wound, we temporarily removed it until it was recovered. This happened on one occasion: α7 was taken out from week 33 to week 36 of the experiment. One rat (α4) was permanently removed from the experiment due to weight loss at week 31 at the age of 37 weeks.

Individual and social interaction tests

These tests were performed at the end of the group measurement period, and included a total of 27 male rats at the age of 44 weeks. Subjects completed a test battery consisting of seven subtests examining fear-related and social behaviors in the following order (see descriptions below): black and white box, canopy, elevated plus-maze, social interaction test with outgroup conspecific, and social interaction test with in-group conspecific. The illumination during these tests was set according to the dark period mentioned above. Body weights were 480±70g (mean±SD) at the time of these tests. The behavioral tests were conducted on three consecutive days between 10 a.m. and 6 p.m. (26 to 28 Feb 2012). The apparatuses were constructed from plastic sheets and cleaned between tests. Depending on the test, either automated analysis was used to obtain trajectories or the behavior was coded by observers using the Solomon Coder software (beta 19.08.02). To ensure inter-observer reliability,

pairs of observers overlapped in 20% of the behavioral tests they scored. Using this overlap, the inter-observer reliability was calculated using the intraclass correlation coefficient (ICC) for all variables except the video frame variables, and we found all observations to be reliable (ICC > 0.9).

1. Black and white box. As described in Ramos et al.[88], the apparatus had a black and white compartment (each sized 27x27x27). The white compartment was strongly illuminated by a white bulb ($\sim$825 lux), while the black compartment was illuminated with a red bulb ($\sim$90 lux; Figure S9A). The bulbs were positioned 37 cm above the apparatus floor. Each rat was initially placed in the center of the white compartment in a direction facing the opposing black compartment, and behavior was then recorded for 5 minutes. We tallied the number of video frames when the animal was in the 1) white compartment, 2) black compartment, or 3) at the border of the two areas and thus could not be clearly assigned (labeled as "both" in the data).

2. Canopy. The apparatus consisted of a circular platform (104 cm in diameter) and a canopy (semitransparent red Plexiglas of 70 cm diameter) 10 cm above the platform (Figure S9B). The mean illumination was 90 lux under the canopy and 400 lux outside of the canopy. At the beginning of the test, the animal was placed under the canopy. The test lasted for 5 minutes. We counted the number of video frames when the animal was 1) under the canopy, 2) in the exposed zone.

3. Elevated plus-maze. Based on Ramos et al.[88], the apparatus had four elevated arms (66 cm from the floor), 45 cm long and 10 cm wide (Figure S9C). Two closed arms enclosed by a 50 cm high wall were located on opposing sides, and two open arms on the other two sides; the wall structure led to different illumination, with 25 lux for the closed arms and 65 lux for the open arms. The central platform (10$\times$ 10 cm) connected the four arms to allow access to any. Each rat was first placed in the central platform facing an open arm, and subsequently, behavior was recorded for 5 minutes. To describe behavior, we counted the number of video frames in which the animal was in the 1) closed arms, 2) open arms, and 3) central platform (labeled as "both" in the data).

4. Social interaction test with an unfamiliar (out-group) conspecific. In an open field arena, we placed an unfamiliar adult male next to a focal rat that had been part of the long-term experiment. Two different unfamiliar rats were used for each phase 3 group (i.e. $G1$ rats were tested with unfamiliar rats 1 and 2, $G2$ rats with unfamiliar rats 3 and 4, etc.). The test apparatus was made out of glass, with a green floor of 74 $\times$ 74 cm and transparent walls ($\sim$40 cm high; Figure S9D). The unfamiliar rats were significantly younger and smaller than the focals (12-weeks-old and 360±20g (mean±SD)). We recorded the behavior for 10 minutes. The test was repeated with other rats (unfamiliar and familiar) after a break that lasted for on average 75±42 min. but a minimum of 35 min. The coded behavior included the following: duration of bipedal orienting stance (%), duration of self-grooming (%), duration of exploration, duration of sniffing non-genital body parts (%), duration of sniffing the genitals of the partner (%), number of steps on the partner, number of fights. The coded parameters were stored separately from the trials with unfamiliar rats 1 and 2 for each focal rat.

5. Social interaction test with a familiar (in-group) conspecific. The social interaction was also performed with a familiar conspecific, chosen as a random groupmate from their phase 3 group. Each individual was measured with four randomly selected members from their group. The coded variables were the same as above, but they were averaged over the trials for each focal rat.

Quantification and statistical analysis

Data processing and behavioral metrics

We calculated metrics of space use from the individual trajectory data for each rat. Time at feeder is the fraction of total time spent at the feeder during nightlight (active period). Distance from wall is the average distance from the walls during nightlight. Home range is the area of an individual's space-use heatmap during all times (number of bins where it was detected more than 10 frames per day, calculated using bins with a linear size of $\sim$ 2 mm) (see Figure 1C), normalized by total number of bins and frames counted. Top of nestbox is the fraction of total time spent on top of the nestbox/shelter area during nightlight.

An approach-avoidance (AA) event was defined for a given pair of individuals ($i \neq j$) if, for a 0.4s long time window, the time-averaged dot product of i's velocity ($v_i(t)$) and the normalized relative direction vector pointing from i to j – a unit vector $\hat{d}_{ij}(t) = d_{ij}(t)/|d_{ij}(t)|$, where $d_{ij}(t) = x_j(t) - x_i(t)$ is the relative position – were within predetermined thresholds for both individuals. The thresholds used were $AA_{ij} = \langle (v_i(t) \cdot \hat{d}_{ji}(t) \rangle_t > 0.8$ for the approacher, and $AA_{ji} < -0.5$ for the avoider. In addition, we used the requirement that i and j were within 40 cm of each other ($|d_{ij}(t)| \leq d_{max} = 40$ cm) and both i and j were moving at speeds of at least 0.25 m/sec ($|v(t)| \geq v_{min} = 0.25$ m/sec).

We use the approach-avoidance event network to quantify the social interaction structure in each group. The values A_{ij} are the number of times rat i dominated approach-avoidance events with rat j. Using this, the number of events rat i dominated is $w_i = \sum_j A_{ij}$, and the number of events lost is $l_i = \sum_j A_{ji}$. The fraction of events dominated for rat i is then

$$f_i = \frac{w_i}{w_i + l_i}. \tag{1}$$

Reaching centrality is calculated using the normalized network of excess wins, W_{ij}. This network has positive entries for

the rat in a pair that dominated in more events and zero for the other rat, and is determined as

$$W_{ij} = \frac{1}{Z} max\left(A_{ij} - A_{ji}, 0\right), \tag{2}$$

where Z is a normalization factor, which we define so that the maximum entry of W_{ij} is equal to 1. This network is then provided as input to the `networkx` function `local_reaching_centrality` to calculate the local reaching centrality (LRC) for each individual, and to the function `global_reaching_centrality` to calculate global reaching centrality (GRC) for the group. Note that we set the flag `normalized=False` for calling these functions, because we use the definition in Eq 2 where W_{ij} is already normalized. With this, the LRC and GRC scores are in the range of 0 to 1.

We used the `EloRating` package[89] to calculate the individual David's score, David's score steepness (DSS), directional consistency index (DCI), and triangle transitivity index (TTRI). For the individual Elo scores, we used the `EloChoice` package[89], which has an improved and more efficient implementation of the randomization of interaction sequences used to calculate the Elo score. We used the `EloSteepness` package[90] to calculate the Elo score steepness (ESS). These *R* packages were integrated into our Python-based analysis code using `rpy2`.

Individual and social interaction assays

We used principal component analysis (PCA) to define the boldness and social interaction scores for each rat from the individual and pairwise assays.

Boldness score. The 8 videoframe variables calculated from the individual tests (black-and-white box, elevated plus-maze, and canopy test) were used to define the boldness score, which reflects the time spent in exposed portions of an unfamiliar environment. We converted each frame count to a fraction of the test time and normalized the input variables before applying PCA. The first component explains 44.7% of the variance, and positive projections onto this component represent more time in open areas (Figure S10A). We used the projection of each rat onto the first component as the "boldness" score. For comparison, we also calculated fractions of open-area time for each test: fraction in the white area during black and white black box test (BWB-whitefrac = BWB-White/(BWB-White + BWB-Black)), fraction of time in open during the elevated cross test (ElevX-openfrac = ElevX-Open/(ElevX-Open + ElevX-Closed)), and fraction of time out during the canopy test (Canopy-outfrac = Canopy-OUT/(Canopy-Out + Canopy-Under)).

Social interaction score. We applied PCA to measures from the pairwise social interaction tests and used results to define a composite score related to social interaction, and additionally used the 2nd PCA component to compare a "self grooming" score. The variables included are duration of sniffing genitals (%), duration of sniffing nongenital body parts (%), duration of bipedal orienting stance (standing up) (%), number of steps on the partner, number of mating attempts, number of fights, duration of exploration (%), and duration of self grooming (%). The components shown in Figure S10B were determined using data from the first test with an unfamiliar rat; the scores for the other tests with a familiar rat and a second unfamiliar rat were calculated by projecting the associated variables onto these components. The first PCA component represents interaction with the other rat, with positive projections indicating more interactions. We use this component as the "social interaction" score. The second PCA component is weighted most strongly by self grooming (positive) and exploration (negative) – we addtionally compare this component as the "self grooming" score.

Figure S13 shows a comparison of the scores, including the boldness score and measures from the individual tests, and the social interaction and self grooming scores from the first test with an unfamiliar rat as well as tests with a familiar rat and a second unfamiliar rat.

References

1. Vishwakarma, M. *et al.* Mechanical interactions among followers determine the emergence of leaders in migrating epithelial cell collectives. *Nat. Commun.* **9**, 3469, DOI: 10.1038/s41467-018-05927-6 (2018).
2. Davidson, J. D., Arauco-Aliaga, R. P., Crow, S., Gordon, D. M. & Goldman, M. S. Effect of interactions between harvester ants on forager decisions. *Front. Ecol. Evol.* **4**, 115, DOI: 10.3389/fevo.2016.00115 (2016).
3. Nagy, M. *et al.* Synergistic Benefits of Group Search in Rats. *Curr. Biol.* **30**, 4733–4738.e4, DOI: 10.1016/j.cub.2020.08.079 (2020).
4. Bond, R. M. *et al.* A 61-million-person experiment in social influence and political mobilization. *Nature* **489**, 295–298, DOI: 10.1038/nature11421 (2012). Number: 7415 Publisher: Nature Publishing Group.
5. Kaburu, S. S. K. & Newton-Fisher, N. E. Egalitarian despots: hierarchy steepness, reciprocity and the grooming-trade model in wild chimpanzees, Pan troglodytes. *Animal Behav.* **99**, 61–71, DOI: 10.1016/j.anbehav.2014.10.018 (2015).
6. Nagy, M. *et al.* Context-dependent hierarchies in pigeons. *Proc. Natl. Acad. Sci.* **110**, 13049–13054, DOI: 10.1073/pnas.1305552110 (2013).

7. Maruska, K. P. Social regulation of reproduction in male cichlid fishes. *Gen. Comp. Endocrinol.* **207**, 2–12, DOI: 10.1016/j.ygcen.2014.04.038 (2014).

8. King, A. J., Douglas, C. M., Huchard, E., Isaac, N. J. & Cowlishaw, G. Dominance and affiliation mediate despotism in a social primate. *Curr. Biol.* **18**, 1833–1838 (2008). Publisher: Elsevier.

9. Stewart, A. J. *et al.* Information gerrymandering and undemocratic decisions. *Nature* **573**, 117–121, DOI: 10.1038/s41586-019-1507-6 (2019). Number: 7772 Publisher: Nature Publishing Group.

10. Majolo, B., Lehmann, J., de Bortoli Vizioli, A. & Schino, G. Fitness-related benefits of dominance in primates. *Am. J. Phys. Anthropol.* **147**, 652–660, DOI: 10.1002/ajpa.22031 (2012). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ajpa.22031.

11. Barnett, S. A. An Analysis of Social Behaviour in Wild Rats. *Proc. Zool. Soc. Lond.* **130**, 107–152, DOI: 10.1111/j.1096-3642.1958.tb00565.x (1958). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1096-3642.1958.tb00565.x.

12. Grant, E. C. An Analysis of the Social Behaviour of the Male Laboratory Rat. *Behaviour* **21**, 260–281, DOI: 10.1163/156853963X00194 (1963). Publisher: Brill.

13. Schweinfurth, M. K. The social life of Norway rats (Rattus norvegicus). *eLife* **9**, e54020, DOI: 10.7554/eLife.54020 (2020). Publisher: eLife Sciences Publications, Ltd.

14. Dennis, E. J. *et al.* Systems Neuroscience of Natural Behaviors in Rodents. *J. Neurosci.* **41**, 911–919 (2021).

15. Forkosh, O. Animal behavior and animal personality from a non-human perspective: Getting help from the machine. *Patterns* **2**, DOI: 10.1016/j.patter.2020.100194 (2021). Publisher: Elsevier.

16. Hinz, R. C. & Polavieja, G. G. d. Ontogeny of collective behavior reveals a simple attraction rule. *Proc. Natl. Acad. Sci.* 201616926, DOI: 10.1073/pnas.1616926114 (2017).

17. Smith, M. L. *et al.* Behavioral variation across the days and lives of honey bees. *iScience* **25**, DOI: 10.1016/j.isci.2022.104842 (2022).

18. Evans, J. C., Lindholm, A. K. & König, B. Long-term overlap of social and genetic structure in free-ranging house mice reveals dynamic seasonal and group size effects. *Curr. Zool.* **67**, 59–69, DOI: 10.1093/cz/zoaa030 (2021).

19. Jabarin, R., Netser, S. & Wagner, S. Beyond the three-chamber test: toward a multimodal and objective assessment of social behavior in rodents. *Mol. Autism* **13**, 41, DOI: 10.1186/s13229-022-00521-6 (2022).

20. Shemesh, Y. *et al.* High-order social interactions in groups of mice. *eLife* **2**, e00759, DOI: 10.7554/eLife.00759 (2013).

21. Forkosh, O. *et al.* Identity domains capture individual differences from across the behavioral repertoire. *Nat. Neurosci.* **22**, 2023–2028, DOI: 10.1038/s41593-019-0516-y (2019). Number: 12 Publisher: Nature Publishing Group.

22. Karamihalev, S. *et al.* Social dominance mediates behavioral adaptation to chronic stress in a sex-specific manner. *eLife* **9**, e58723, DOI: 10.7554/eLife.58723 (2020). Publisher: eLife Sciences Publications, Ltd.

23. Lopez, J. P. *et al.* Ketamine exerts its sustained antidepressant effects via cell-type-specific regulation of Kcnq2. *Neuron* **110**, 2283–2298.e9, DOI: 10.1016/j.neuron.2022.05.001 (2022).

24. Williamson, C. M., Lee, W. & Curley, J. P. Temporal dynamics of social hierarchy formation and maintenance in male mice. *Animal Behav.* **115**, 259–272, DOI: 10.1016/j.anbehav.2016.03.004 (2016).

25. Lee, W. *et al.* Effect of relative social rank within a social hierarchy on neural activation in response to familiar or unfamiliar social signals. *Sci. Reports* **11**, 2864, DOI: 10.1038/s41598-021-82255-8 (2021). Number: 1 Publisher: Nature Publishing Group.

26. Lee, W. *et al.* Distinct immune and transcriptomic profiles in dominant versus subordinate males in mouse social hierarchies. *Brain, Behav. Immun.* **103**, 130–144, DOI: 10.1016/j.bbi.2022.04.015 (2022).

27. Johnson, J. A. Supplanting by olive baboons: dominance rank difference and resource value. *Behav. Ecol. Sociobiol.* **24**, 277–283, DOI: 10.1007/BF00290903 (1989).

28. Evans, J. C., Devost, I., Jones, T. B. & Morand-Ferron, J. Inferring dominance interactions from automatically recorded temporal data. *Ethology* **124**, 188–195, DOI: 10.1111/eth.12720 (2018).

29. Gullstrand, J., Claidière, N. & Fagot, J. Computerized assessment of dominance hierarchy in baboons (Papio papio). *Behav. Res. Methods* **53**, 1923–1934, DOI: 10.3758/s13428-021-01539-z (2021).

30. Spruijt, B. M., Hol, T. & Rousseau, J. Approach, avoidance, and contact behavior of individually recognized animals automatically quantified with an imaging technique. *Physiol. & Behav.* **51**, 747–752, DOI: 10.1016/0031-9384(92)90111-E (1992).

31. Elo, A. E. *The rating of chessplayers, past and present* (Arco Pub, New York, 1978).

32. Albers, P. C. & De Vries, H. Elo-rating as a tool in the sequential estimation of dominance strengths. *Animal Behav.* **61**, 489–495, DOI: 10.1006/anbe.2000.1571 (2001).

33. Neumann, C. *et al.* Assessing dominance hierarchies: validation and advantages of progressive evaluation with Elo-rating. *Animal Behav.* **82**, 911–921, DOI: 10.1016/j.anbehav.2011.07.016 (2011).

34. Strauss, E. D. & Holekamp, K. E. Inferring longitudinal hierarchies: Framework and methods for studying the dynamics of dominance. *J. Animal Ecol.* **88**, 521–536, DOI: 10.1111/1365-2656.12951 (2019). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1365-2656.12951.

35. Mones, E., Vicsek, L. & Vicsek, T. Hierarchy Measure for Complex Networks. *PLOS ONE* **7**, e33799, DOI: 10.1371/journal.pone.0033799 (2012). Publisher: Public Library of Science.

36. Ozogány, K. & Vicsek, T. Modeling the Emergence of Modular Leadership Hierarchy During the Collective Motion of Herds Made of Harems. *J. Stat. Phys.* **158**, 628–646, DOI: 10.1007/s10955-014-1131-7 (2015).

37. Shimoji, H., Abe, M. S., Tsuji, K. & Masuda, N. Global network structure of dominance hierarchy of ant workers. *J. The Royal Soc. Interface* **11**, 20140599, DOI: 10.1098/rsif.2014.0599 (2014). Publisher: Royal Society.

38. Kora, Y., Salhi, S., Davidsen, J. & Simon, C. Global excitability and network structure in the human brain. *Phys. Rev. E* **107**, 054308, DOI: 10.1103/PhysRevE.107.054308 (2023).

39. Hu, F., Zhao, S., Bing, T. & Chang, Y. Hierarchy in industrial structure: The cases of China and the USA. *Phys. A: Stat. Mech. its Appl.* **469**, 871–882, DOI: 10.1016/j.physa.2016.11.083 (2017).

40. Beardsley, K., Liu, H., Mucha, P. J., Siegel, D. A. & Tellez, J. F. Hierarchy and the Provision of Order in International Politics. *The J. Polit.* **82**, 731–746, DOI: 10.1086/707096 (2020). Publisher: The University of Chicago Press.

41. Mones, E., Pollner, P. & Vicsek, T. Universal hierarchical behavior of citation networks. *J. Stat. Mech. Theory Exp.* **2014**, P05023, DOI: 10.1088/1742-5468/2014/05/P05023 (2014). Publisher: IOP Publishing and SISSA.

42. Strauss, E. D. *et al.* DomArchive: a century of published dominance data. *Philos. Transactions Royal Soc. B: Biol. Sci.* **377**, 20200436, DOI: 10.1098/rstb.2020.0436 (2022). Publisher: Royal Society.

43. Shizuka, D. & McDonald, D. B. A social network perspective on measurements of dominance hierarchies. *Animal Behav.* **83**, 925–934, DOI: 10.1016/j.anbehav.2012.01.011 (2012).

44. de Vries, H. An improved test of linearity in dominance hierarchies containing unknown or tied relationships. *Animal Behav.* **50**, 1375–1389, DOI: 10.1016/0003-3472(95)80053-0 (1995).

45. Neumann, C. & Fischer, J. Extending Bayesian Elo-rating to quantify the steepness of dominance hierarchies. *Methods Ecol. Evol.* **14**, 669–682, DOI: 10.1111/2041-210X.14021 (2023). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.14021.

46. Burt, W. H. Territoriality and home range concepts as applied to mammals. *J. mammalogy* **24**, 346–352 (1943).

47. Varholick, J. A. *et al.* Social dominance hierarchy type and rank contribute to phenotypic variation within cages of laboratory mice. *Sci. Reports* **9**, 13650, DOI: 10.1038/s41598-019-49612-0 (2019). Publisher: Nature Publishing Group.

48. de Vries, H., Stevens, J. M. G. & Vervaecke, H. Measuring and testing the steepness of dominance hierarchies. *Animal Behav.* **71**, 585–592, DOI: 10.1016/j.anbehav.2005.05.015 (2006).

49. van Hooff, J. A. R. A. M. & Wensing, J. A. B. Dominance and its behavioral measures in a captive wolf pack. In *Man and wolf: Advances, issues, and problems in captive wolf research*, Perspectives in vertebrate science, Vol. 4., 219–252 (Dr W Junk Publishers, Dordrecht, Netherlands, 1987).

50. Shizuka, D. & McDonald, D. B. The network motif architecture of dominance hierarchies. *J. The Royal Soc. Interface* **12**, 20150080, DOI: 10.1098/rsif.2015.0080 (2015). Publisher: Royal Society.

51. Boreman, J. & Price, E. Social dominance in wild and domestic Norway rats (Rattus norvegicus). *Animal Behav.* **20**, 534–542, DOI: 10.1016/S0003-3472(72)80018-6 (1972).

52. Macdonald, D. W., Berdoy, M. & Smith, P. Stability of Social Status in Wild Rats: Age and the Role of Settled Dominance. *Behaviour* **132**, 193–212, DOI: 10.1163/156853995X00694 (1995). Publisher: Brill.

53. File, S. E. & Hyde, J. R. Can social interaction be used to measure anxiety? *Br. J. Pharmacol.* **62**, 19–24 (1978).

54. File, S. E. & Seth, P. A review of 25 years of the social interaction test. *Eur. J. Pharmacol.* **463**, 35–53, DOI: 10.1016/S0014-2999(03)01273-1 (2003).

55. Bolivar, V. J., Walters, S. R. & Phoenix, J. L. Assessing autism-like behavior in mice: Variations in social interactions among inbred strains. *Behav. Brain Res.* **176**, 21–26, DOI: 10.1016/j.bbr.2006.09.007 (2007).

56. Silverman, J. L., Yang, M., Lord, C. & Crawley, J. N. Behavioural phenotyping assays for mouse models of autism. *Nat. Rev. Neurosci.* **11**, 490–502, DOI: 10.1038/nrn2851 (2010). Number: 7 Publisher: Nature Publishing Group.

57. Calhoun, J. B. *The Ecology and Sociology of the Norway Rat* (U.S. Department of Health, Education, and Welfare, Public Health Service, 1963). Google-Books-ID: crhqAAAAMAAJ.

58. Bolles, R. C. & Woods, P. J. The ontogeny of behaviour in the albino rat. *Animal Behav.* **12**, 427–441, DOI: 10.1016/0003-3472(64)90062-4 (1964).

59. Spink, A. J., Tegelenbosch, R. A. J., Buma, M. O. S. & Noldus, L. P. J. J. The EthoVision video tracking system—A tool for behavioral phenotyping of transgenic mice. *Physiol. & Behav.* **73**, 731–744, DOI: 10.1016/S0031-9384(01)00530-3 (2001).

60. Hong, W. *et al.* Automated measurement of mouse social behaviors using depth sensing, video tracking, and machine learning. *Proc. Natl. Acad. Sci.* **112**, E5351–E5360, DOI: 10.1073/pnas.1515982112 (2015). Publisher: Proceedings of the National Academy of Sciences.

61. de Chaumont, F. *et al.* Real-time analysis of the behaviour of groups of mice via a depth-sensing camera and machine learning. *Nat. Biomed. Eng.* **3**, 930–942, DOI: 10.1038/s41551-019-0396-1 (2019). Number: 11 Publisher: Nature Publishing Group.

62. Nilsson, S. R. *et al.* Simple Behavioral Analysis (SimBA) – an open source toolkit for computer classification of complex social behaviors in experimental animals, DOI: 10.1101/2020.04.19.049452 (2020). Pages: 2020.04.19.049452 Section: New Results.

63. Fong, T., Jury, B., Hu, H. & Murphy, T. H. PyRodentTracks: flexible computer vision and RFID based system for multiple rodent tracking and behavioral assessment, DOI: 10.1101/2022.01.23.477395 (2022). Pages: 2022.01.23.477395 Section: New Results.

64. Szechtman, H., Dvorkin-Gheva, A. & Gomez-Marin, A. A virtual library for behavioral performance in standard conditions—rodent spontaneous activity in an open field during repeated testing and after treatment with drugs or brain lesions. *GigaScience* **11**, giac092, DOI: 10.1093/gigascience/giac092 (2022).

65. Schweinfurth, M. K., Stieger, B. & Taborsky, M. Experimental evidence for reciprocity in allogrooming among wild-type Norway rats. *Sci. Reports* **7**, 4010, DOI: 10.1038/s41598-017-03841-3 (2017). Number: 1 Publisher: Nature Publishing Group.

66. Kim, D. G. *et al.* Social Interaction Test in Home Cage as a Novel and Ethological Measure of Social Behavior in Mice. *Exp. Neurobiol.* **28**, 247–260, DOI: 10.5607/en.2019.28.2.247 (2019).

67. Acikgoz, B., Dalkiran, B. & Dayi, A. An overview of the currency and usefulness of behavioral tests used from past to present to assess anxiety, social behavior and depression in rats and mice. *Behav. Process.* **200**, 104670, DOI: 10.1016/j.beproc.2022.104670 (2022).

68. Puścian, A. & Knapska, E. Blueprints for measuring natural behavior. *iScience* **25**, DOI: 10.1016/j.isci.2022.104635 (2022). Publisher: Elsevier.

69. Pearce, J. M. *Animal Learning and Cognition: An Introduction* (Psychology Press, London, 2014), 3 edn.

70. Phifer-Rixey, M. & Nachman, M. W. Insights into mammalian biology from the wild house mouse Mus musculus. *eLife* **4**, e05959, DOI: 10.7554/eLife.05959 (2015). Publisher: eLife Sciences Publications, Ltd.

71. Bale, T. L. *et al.* The critical importance of basic animal research for neuropsychiatric disorders. *Neuropsychopharmacology* **44**, 1349–1353, DOI: 10.1038/s41386-019-0405-9 (2019). Number: 8 Publisher: Nature Publishing Group.

72. Homberg, J. R. Measuring behaviour in rodents: towards translational neuropsychiatric research. *Behav. brain research* **236**, 295–306 (2013).

73. Peters, S. M., Pothuizen, H. H. J. & Spruijt, B. M. Ethological concepts enhance the translational value of animal models. *Eur. J. Pharmacol.* **759**, 42–50, DOI: 10.1016/j.ejphar.2015.03.043 (2015).

74. Kondrakiewicz, K., Kostecki, M., Szadzińska, W. & Knapska, E. Ecological validity of social interaction tests in rats and mice. *Genes, Brain Behav.* **18**, e12525, DOI: 10.1111/gbb.12525 (2019). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gbb.12525.

75. Shemesh, Y. & Chen, A. A paradigm shift in translational psychiatry through rodent neuroethology. *Mol. Psychiatry* 1–11, DOI: 10.1038/s41380-022-01913-z (2023). Publisher: Nature Publishing Group.

76. Grieco, F. *et al.* Measuring Behavior in the Home Cage: Study Design, Applications, Challenges, and Perspectives. *Front. Behav. Neurosci.* **15**, 735387, DOI: 10.3389/fnbeh.2021.735387 (2021).

77. Gunaydin, L. A. *et al.* Natural neural projection dynamics underlying social behavior. *Cell* **157**, 1535–1551 (2014).

78. Padilla-Coreano, N., Tye, K. M. & Zelikowsky, M. Dynamic influences on the neural encoding of social valence. *Nat. Rev. Neurosci.* **23**, 535–550 (2022).

79. Lee, C. R., Chen, A. & Tye, K. M. The neural circuitry of social homeostasis: Consequences of acute versus chronic social isolation. *Cell* **184**, 1500–1516 (2021).

80. Mathis, A. *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289, DOI: 10.1038/s41593-018-0209-y (2018). Number: 9 Publisher: Nature Publishing Group.

81. Graving, J. M. *et al.* DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. *eLife* **8**, e47994, DOI: 10.7554/eLife.47994 (2019).

82. Pereira, T. D. *et al.* Fast animal pose estimation using deep neural networks. *Nat. Methods* **16**, 117–125, DOI: 10.1038/s41592-018-0234-5 (2019). Number: 1 Publisher: Nature Publishing Group.

83. Pereira, T. D. *et al.* SLEAP: A deep learning system for multi-animal pose tracking. *Nat. Methods* **19**, 486–495, DOI: 10.1038/s41592-022-01426-1 (2022). Number: 4 Publisher: Nature Publishing Group.

84. Lehner, S. R., Rutte, C. & Taborsky, M. Rats Benefit from Winner and Loser Effects. *Ethology* **117**, 949–960, DOI: 10.1111/j.1439-0310.2011.01962.x (2011). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1439-0310.2011.01962.x.

85. Pellis, S. *et al.* Measuring play fighting in rats: a multilayered approach. *Curr. Protoc.* **2**, e337 (2022).

86. Ham, J. R. & Pellis, S. M. The goldilocks principle: Balancing familiarity and novelty in the selection of play partners in groups of juvenile male rats. *Animal Behav. Cogn.* **10**, 304–328 (2023).

87. Fulenwider, H. D., Caruso, M. A. & Ryabinin, A. E. Manifestations of domination: Assessments of social dominance in rodents. *Genes, Brain Behav.* **21**, e12731 (2022).

88. Ramos, A., Berton, O., Mormède, P. & Chaouloff, F. A multiple-test study of anxiety-related behaviours in six inbred rat strains. *Behav. brain research* **85**, 57–69 (1997).

89. Neumann, C. EloChoice: Preference Rating for Visual Stimuli Based on Elo Ratings (2019).

90. Neumann, C. EloSteepness: Bayesian Dominance Hierarchy Steepness via Elo Rating and David's Scores (2023).